SCIENCE CHINA
Life Sciences

• RESEARCH PAPER •

July 2013 Vol.56 No.7: 638–646
doi: 10.1007/s11427-013-4497-x

Secretory/releasing proteome-based identification of plasma biomarkers in HBV-associated hepatocellular carcinoma

YANG Lei[1], RONG WeiQi[2], XIAO Ting[1], ZHANG Ying[1], XU Bin[3], LIU Yu[1], WANG LiMing[2], WU Fan[2], QI Jun[4] , ZHAO XiuYing[5], WANG HongXia[3], HAN NaiJun[1], GUO SuPing[1], WU JianXiong[2*], GAO YanNing[1*] & CHENG ShuJun[1*]

[1]*State Key Laboratory of Molecular Oncology, Beijing Key Laboratory for Carcinogenesis and Cancer Prevention, Cancer Institute (Hospital), Peking Union Medical College & Chinese Academy of Medical Sciences, Beijing 100021, China;*
[2]*Department of Abdominal Surgery, Cancer Institute (Hospital), Peking Union Medical College & Chinese Academy of Medical Sciences, Beijing 100021, China;*
[3]*National Center of Biomedical Analysis, Beijing 100850, China;*
[4]*Clinical Laboratory, Cancer Institute (Hospital), Peking Union Medical College & Chinese Academy of Medical Sciences, Beijing 100021, China;*
[5]*Clinical Laboratory, Beijing You'an Hospital, Capital Medical University, Beijing 100069, China*

Received April 3, 2013; accepted May 17, 2013; published online June 7, 2013

For successful therapy, hepatocellular carcinoma (HCC) must be detected at an early stage. Herein, we used a proteomic approach to analyze the secretory/releasing proteome of HCC tissues to identify plasma biomarkers. Serum-free conditioned media (CM) were collected from primary cultures of cancerous tissues and surrounding noncancerous tissues. Proteomic analysis of the CM proteins permitted the identification of 1365 proteins. The enriched molecular functions and biological processes of the CM proteins, such as hydrolase activity and catabolic processes, were consistent with the liver being the most important metabolic organ. Moreover, 19% of the proteins were characterized as extracellular or membrane-bound. For validation, secretory proteins involved in transforming growth factor-β signaling pathways were validated in plasma samples. Alpha-fetoprotein (AFP), metalloproteinase (MMP)1, osteopontin (OPN), and pregnancy-specific beta-1-glycoprotein (PSG)9 were significantly increased in HCC patients. The overall performance of MMP1 and OPN in the diagnosis of HCC remained greater than that of AFP. In addition, this study represents the first report of MMP1 as a biomarker with a higher sensitivity and specificity than AFP. Thus, this study provides a valuable resource of the HCC secretome with the potential to investigate serological biomarkers. MMP1 and OPN could be used as novel biomarkers for the early detection of HCC and to improve the sensitivity of biomarkers compared with AFP.

hepatocellular carcinoma (HCC), secretome, biomarker, MMP1, OPN, PSG9

Citation: Yang L, Rong W Q, Xiao T, et al. Secretory/releasing proteome-based identification of plasma biomarkers in HBV-associated hepatocellular carcinoma. Sci China Life Sci, 2013, 56: 638–646, doi: 10.1007/s11427-013-4497-x

Approximately 600000 new cases of hepatocellular carcinoma (HCC) are diagnosed each year, of which approximately 55% are detected in China [1]. The relationship between HCC and persistent hepatitis B virus (HBV) infection has been well documented. More than 20 million individuals suffer from chronic HBV infection in China [2]. Surgical resection is the most effective treatment for HCC patients, but the reported five-year disease-free survival rate is only

*Corresponding author (email: chengshj@263.net.cn; yngao@cicams.ac.cn; dr_wujianxiong@yahoo.com.cn)

25%–30% [3–4]. The majority of HCC diagnoses usually occur at a late stage, resulting in poor prognoses. Alpha-fetoprotein (AFP) has been widely used for high-risk population screening, in combination with ultrasound imaging. However, its limited performance and sensitivity of 39%–65% are not sufficient for accurate screening [5]. Thus, improvements in the early detection and more accurate markers are urgently needed.

Proteomics analysis has provided critical and specific information regarding the functional and structural activities of proteins in organisms. One major challenge in plasma proteomics is the large dynamic range of protein abundances, from mg mL^{-1} to pg mL^{-1} or lower [6]. A small number of high-abundance proteins mask the majority of low-abundance proteins. To address this problem, several approaches have used affinity-based depletion, ultracentrifugation, or precipitation to remove the most abundant proteins from samples. Unfortunately, these methods have low reproducibility, low specificity, or low-throughput and are thus not suitable for identifying novel proteins [7]. Because of the formidable challenges of direct analysis of plasma/serum samples, the secretome has been pursued as an alternative approach for blood biomarker discovery. Secreted proteins are also most likely excellent candidate serological tumor markers because they are released by cells and thus have the highest potential to enter the circulation [8]. For example, Planque et al. [9] identified five biomarker candidates in plasma by analyzing conditioned media (CM) from four lung cancer cell lines. Polisetty et al. [10] constructed a secretome of three glioblastoma cell lines and demonstrated that 22 proteins were detectable in plasma. However, few reports have focused on the HCC secretome for screening biomarkers in blood.

Cell-based systems for proteomics analysis have emerged as a time- and cost-effective platform for identifying biomarkers. Nevertheless, the discrepancy between immortalized cancer cells and clinical samples may be considerable. Therefore, we previously developed a novel approach based on the primary culture of tumor tissues and proteomic analysis of serum-free CM [11,12]. In this study, we collected serum-free CM from cultures of cancerous and surrounding noncancerous tissues from six HCC patients with HBV infection. Proteins in the CM were investigated by MS-based proteomic analysis to establish a high-quality secretory/releasing proteome of HCC. We then validated AFP and biomarker candidates including matrix metalloproteinase (MMP)1, MMP7, MMP9, osteopontin (OPN), pregnancy-specific beta-1-glycoprotein (PSG)9, and tissue inhibitor of metalloproteinase (TIMP)1 in human plasma and evaluated their performance for the diagnosis of HCC.

1 Materials and methods

1.1 Patients and specimens

Primary tissue cultures, including cancerous and surrounding noncancerous tissue samples, were obtained from six primary HCC patients. All patients were HBV carriers who underwent surgical resection in March 2009 at the Cancer Hospital, Peking Union Medical College (PUMC), Chinese Academy of Medical Sciences (CAMS). The histological diagnosis of the tissue samples was confirmed by experienced pathologists. The corresponding clinicopathological factors are described in Table S1 in Supporting Information.

HCC plasma was collected from 179 HCC patients (35 females and 144 males, with a mean age of 54.0 years) before hepatectomy at the Cancer Hospital, PUMC & CAMS. All HCC patients were chronically infected with HBV. Cirrhosis plasma was obtained from 80 liver cirrhosis patients (24 females and 56 males, with a mean age of 53.5 years) with chronic HBV infection at Beijing You'an Hospital, Capital Medical University. Healthy plasma from 103 healthy controls without HBV infection (13 females and 90 males, with a mean age of 49.1 years) was obtained from a health screening program at the Cancer Hospital. Peripheral blood was obtained as described previously [12]. All tissue and plasma samples were collected with informed consent and approval of the Institute Ethics Committee.

1.2 Primary organ culture and conditioned medium harvest

Primary organ cultures were based on six paired tissues containing cancerous and surrounding noncancerous tissues from HCC patients subjected to hepatectomy. The freshly obtained tissue samples were rinsed, cut into 2–3 mm^3 pieces, and transferred to 60 mm culture dishes. The explants were cultivated at 37°C in LHC-9 medium (serum-free) [13] in a gas mixture of 50% O_2, 45% N_2, and 5% CO_2. Hank's solution was used to wash the cultures 24 h later. The explants were then cultivated in serum-free LHC-9 without bovine pituitary extract for another 24 h. The CM was then collected and dialyzed with an Amicon Ultra-15 centrifugal filter device (Millipore, Billerica, MA, USA).

1.3 One-dimensional sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and liquid chromatography-mass spectrometry (LC-MS)/MS analysis of CM proteins

CM samples containing 50 μg of total protein were subjected to 10% SDS-PAGE. After staining with Coomassie blue, 40 bands were excised from the gel lanes and trypsin-digested according to Zhang et al. [12]. Then, nano-LC-MS/MS was conducted with a SYNAPT system (Waters, MA, USA). Each peptide mixture extract was dissolved in

6 μL 0.1% formic acid and separated on a nano-ACQUITY system (Waters) equipped with a Symmetry C18 pre-column and an ethylene bridged hybrid (EBH) C18 analytical reversed-phase column. The samples were transferred from an aqueous 0.1% formic acid solution to the pre-column at a rate of 7 μL min^{-1} for 3 min. The solvents used were aqueous 0.1% formic acid (v/v) (A) and 0.1% formic acid (v/v) in acetonitrile (B) with a linear gradient of 97%–60% mobile phase A in mobile phase B over 90 min at 200 nL min^{-1}, followed by 10 min of 10% mobile phase A and 90% mobile phase B. The lock mass was delivered to the nano-ACQUITY auxiliary pump with a constant flow of 300 nL min^{-1} at a concentration of 100 fmol μL^{-1} Glu-fibrinopeptide B. Then, the tryptic peptides were analyzed on a SYNAPT high-definition mass spectrometer [14]. The mass spectrometer was operated in the v-mode, at least 10000 full-width half maximum. The TOF analyzer of the mass spectrometer was calibrated with MS/MS fragment ions of Glu-fibrinopeptide B from *m*/*z* 50 to 1600. The sprayer was sampled with a frequency of 30 s. Accurate mass LC-MS data were collected in the high-definition MSE mode (low collision energy 4 eV, high collision energy, ramping from 15 to 45 eV, switching every 1 s, with an interscan time of 0.02 s). The mass range was from 300 to 1990 *m*/*z*.

1.4 CM protein identification

The protein-Lynx Global Server version 2.3 (Waters) was used to generate peak list files for the SYNAPT spectra. Those files were then searched against UniProt (human, 76137 entries; UniProt Knowledgebase Release 12.6 consists of UniProtKB/Swiss-Prot Release 54.6 from December 4, 2007) using the pFind 2.4 search program. In all database searches, trypsin was designated as the protease, and only one missed cleavage was allowed. The maximal mass tolerance was set at 0.2 Da for the precursor ions and 0.3 Da for the fragment ions. Carbamidomethylation (Cysteine) was searched as a fixed modification. Oxidation (methionine) was set as a variable modification. The overall number of false positives can be estimated by doubling the number of peptides found from the reverse sequences. A peptide false discovery rate (FDR) lower than 1% was applied. The pFind 2.4 program provides a fully automated method. Finally, the remaining peptides were identified according to the protein inference algorithm .

1.5 Measurement of the CM proteins in plasma

All experimental procedures were performed according to the manufacturer's instructions. AFP levels were tested using a commercial immunoassay with enhanced chemiluminescence at the Clinical Diagnostic Laboratories of the Cancer Hospital, Chinese Academy of Medical Sciences. The levels of six other selected CM proteins (MMP1, MMP7, MMP9, OPN, PSG9 and TIMP1) were evaluated by double-antibody sandwich enzyme-linked immunosorbent assay (ELISA). All ELISA kits were obtained from commercial manufacturers (R&D, Minneapolis, MN, USA) except PSG9 (Blue Gene, Shanghai, China).

1.6 Bioinformatics analysis

Gene ontology (GO) enrichment was performed using GOFFA tools [15]. The ArrayTrack Pathway tool was applied to extract the biological associations of the CM proteins using public databases, including the KEGG and PATHART databases [15]. The CM protein networks were analyzed using STRING online tools. The Mann-Whitney *U*-test (SPSS version 17.0, Chicago, IL, USA) was used to compare the differences in the plasma levels of selected proteins. Receiver operating characteristic (ROC) curves and the area under the ROC (AUC) calculations were performed with SPSS. A *P*-value less than 0.05 was considered statistically significant. Box and whisker plots of the data were prepared by GraphPad Prism for Windows, version 5.0 (GraphPad Software, San Diego, CA, USA).

2 Results

2.1 Establishment of the HCC-related CM protein database

One-dimensional SDS-PAGE was used to separate the CM proteins from the cultured HCC tissues. LC-MS/MS analysis and the pFind search program identified CM proteins from the 12 samples separately as described in Materials and methods. The mean number of identified proteins was 511, with a range of 271 to 772. Figure S1A shows the number of overlapping proteins between the paired tissue CM samples. We identified 1365 proteins, with 841 (61.61%) overlapping proteins and 417 (30.55%) and 107 (7.84%) differential proteins in the cancerous and noncancerous CM samples, respectively (Figure S1A). File S1 in Supporting Information lists detailed information for all proteins identified. In addition, we investigated the molecular weights and pI (isoelectric point) values of the proteins. The molecular weights ranged from 6584 to 669735 Da and were predominantly in the 10–40 kD range (Figure S2A). The pI values ranged from 0.99 to 11.85 and were mainly in the 4.0–6.6 range, with a valley around 5.6 (Figure S2B). Intriguingly, this pattern coincided with the predicted proteomes of different organisms and implied biological activity typical of an acid microenvironment in tumor tissues [16].

We compared two plasma proteome datasets with the HCC CM proteins: the proteins of the HUPO Plasma Proteome Project [17] and the human plasma proteins identified by Anderson et al. [18]. As shown in Figure S1B, 301 and 173 CM proteins were extracted from the HUPO Plasma

Proteome and Anderson's plasma database, respectively. In addition, the HUPO liver proteomic database, which included 9020 proteins [19], was compared, which revealed that 1077 CM proteins were shared with the HUPO liver database. However, 217 proteins in our dataset were not present in these three published datasets (Figure S1B).

2.2 Gene ontology annotation, assignments of pathways and network construction for the CM proteins

The CM proteins were assigned enriched molecular functions, biological processes, and cellular components based on GO enrichment. The enriched molecular functions and biological processes indicated that the liver was the most important metabolic organ. A majority of CM proteins were involved in the following molecular functions: hydrolase activity (26%), enzyme regulation (9%), and lipid binding (6%) (Figure 1A). Biological processes involving the CM proteins mainly included protein metabolism (19%), catabolic processes (17%), and amino acid metabolism (6%) (Figure 1B). With respect to the cellular distributions of the CM proteins, extracellular and membrane proteins accounted for 19% of the total proteins, which are most likely secreted into the blood. Many CM proteins were located in the cytosol, representing 21% of the total proteins. The remaining proteins were distributed in the mitochondria, nucleus, cytoskeleton, Golgi apparatus, and other organelle components (Figure 1C). To understand the nature of the CM proteins, the preferred pathways of the extracellular proteins were classified using the KEGG and PATHART pathways databases. Extracellular proteins participated in pathways including cell communication, immune system, and cell signaling. Furthermore, transforming growth factor (TGF)-β signaling is disrupted in HCC [20]. The interaction networks for CM proteins involved in the TGF-β signaling pathway were investigated by STRING (Figure S3). Sub-network analysis revealed that CM proteins, such as MMP1, MMP9, OPN, and TIMP1 interacted with the TGF-β-induced (TGFBI) protein. In addition, AFP combined with PSG9 participated in the TGF-β signaling pathway by indirectly interacting with other CM proteins.

2.3 Detection of selected CM proteins in plasma

Seven proteins (AFP, MMP1, MMP7, MMP9, OPN, PSG9, and TIMP1) were selected based on the TGF-β signaling pathway and network analyses. To evaluate the concentration of CM proteins in plasma, three cohorts including 179 samples from HCC individuals with HBV infection, 80 samples from individuals with liver cirrhosis and HBV infection, and 103 samples from healthy controls were enrolled. The plasma levels of AFP were significantly higher in the HCC patients (Figure 2A, Table 1). In general, MMP1, OPN, and PSG9 gradually increased in the plasma of healthy controls, liver cirrhosis patients, and HCC patients. MMP1 and OPN were significantly higher in HCC patients than in the healthy controls and liver cirrhosis patients (Figure 2B–D). MMP7, MMP9, and TIMP1 were increased in the plasma of liver cirrhosis patients compared with the healthy controls and HCC patients (Table 1).

2.4 Performance of biomarker candidates

MMP1, OPN, and PSG9 were elevated in HCC samples and thus may be potential plasma biomarker candidates. ROC analysis evaluated the performance of MMP1, OPN, and PSG9. Table 2 shows that the AUCs of MMP1 and OPN but not PSG9 had better performance than AFP in discriminating HCC from healthy controls or cirrhosis (Table 2). When HCC was compared with cirrhosis, the AUC for AFP decreased, but MMP1 had an even higher AUC (0.945; 95% confidence interval (CI): 0.910–0.979) (Figure 3A). When only HCC samples with AFP levels less than 20 ng mL^{-1} were included, the MMP1 AUC remained high in discerning HCC from liver cirrhosis (0.948; 95% CI: 0.912–0.984) (Figure 3B). Moreover, the AUC of OPN (0.911, 95% CI: 0.866–0.957) was greater than that of AFP in discerning HCC from liver cirrhosis (Figure 3C). The AUC of OPN also remained greater than that of AFP for HCC with AFP levels less than 20 ng mL^{-1} (Figure 3D).

When using the currently recommended clinical cutoff for AFP (20 ng mL^{-1}) and the best cutoffs for MMP1 (1.72 ng mL^{-1}) and OPN (84.40 ng mL^{-1}), as determined using

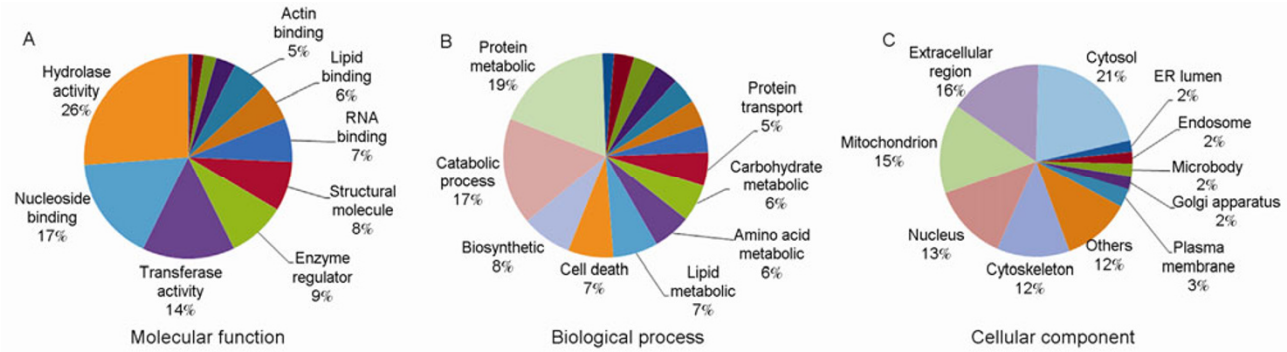

Figure 1 Classification of HCC CM proteins using the gene ontology enrichment tool. A, Molecular function. B, Biological process. C, Cellular component.

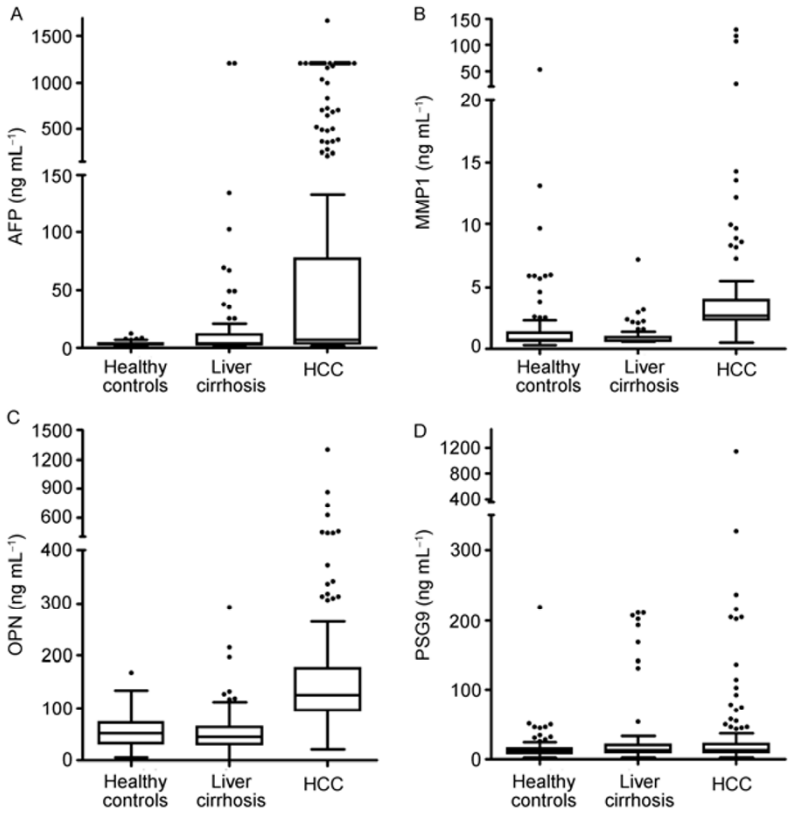

Figure 2 Plasma levels of AFP, MMP1, OPN, and PSG9 in HCC patients, liver cirrhosis patients and healthy controls. A, AFP levels were higher in HCC than in liver cirrhosis and healthy controls. B, MMP1 levels were higher in HCC than in liver cirrhosis and healthy controls. C, OPN plasma levels were increased in HCC compared with the other groups. D, PSG9 plasma levels were elevated in HCC compared with healthy controls. The box refers to the 25th and 75th percentile values, with the line indicating the median values. Points outside the interquartile range are outliers.

Table 1 Mann-Whitney *U*-test analysis levels of selected proteins in the plasma of HCC patients, liver cirrhosis patients, and healthy controls

Protein biomarkers	Median and standard deviation (ng mL^{-1})			*P*-value
	Healthy controls (*n*=103)	Liver cirrhosis (*n*=80)	HCC (*n*=179)	
AFP	2.84 (1.78)	3.84 (109.69)	6.45 (401.62)	<0.001[a)], 0.001[b)], 0.004[c)]
MMP1	0.76 (5.46)	0.63 (0.89)	2.67 (15.00)	<0.001[a)], <0.001[b)], 0.326[c)]
MMP7	1.52 (0.70)	15.11 (11.49)	1.45 (10.20)	0.732[a)], <0.001[b)], <0.001[c)]
MMP9	44.45 (28.19)	150.74 (132.50)	66.30 (49.67)	<0.001[a)], <0.001[b)], <0.001[c)]
OPN	52.20 (30.11)	45.09 (46.59)	125.09 (138.91)	<0.001[a)], <0.001[b)], 0.353[c)]
PSG9	10.68 (22.49)	12.61 (52.57)	12.97 (94.15)	0.005[a)], 0.671[b)], 0.056[c)]
TIMP1	71.62 (14.63)	225.58 (269.58)	128.72 (77.39)	<0.001[a)], <0.001[b)], <0.001[c)]

a) Healthy controls versus HCC. b) Liver cirrhosis versus HCC. c) Liver cirrhosis versus healthy controls.

the minimal distance to the top-left corner in the ROC curve, MMP1 had the best performance, with a sensitivity of 91.62% (95% CI: 86.60–95.20) and a specificity of 91.25% (95% CI: 82.80–96.40) in differentiating HCC from liver cirrhosis (Table 2). For HCC with AFP levels less than 20 ng mL^{-1}, the sensitivity of MMP1 increased to 92.92% (95% CI: 86.50–96.90) with a specificity of 91.25% (95% CI: 82.80–96.64) (Table 2).

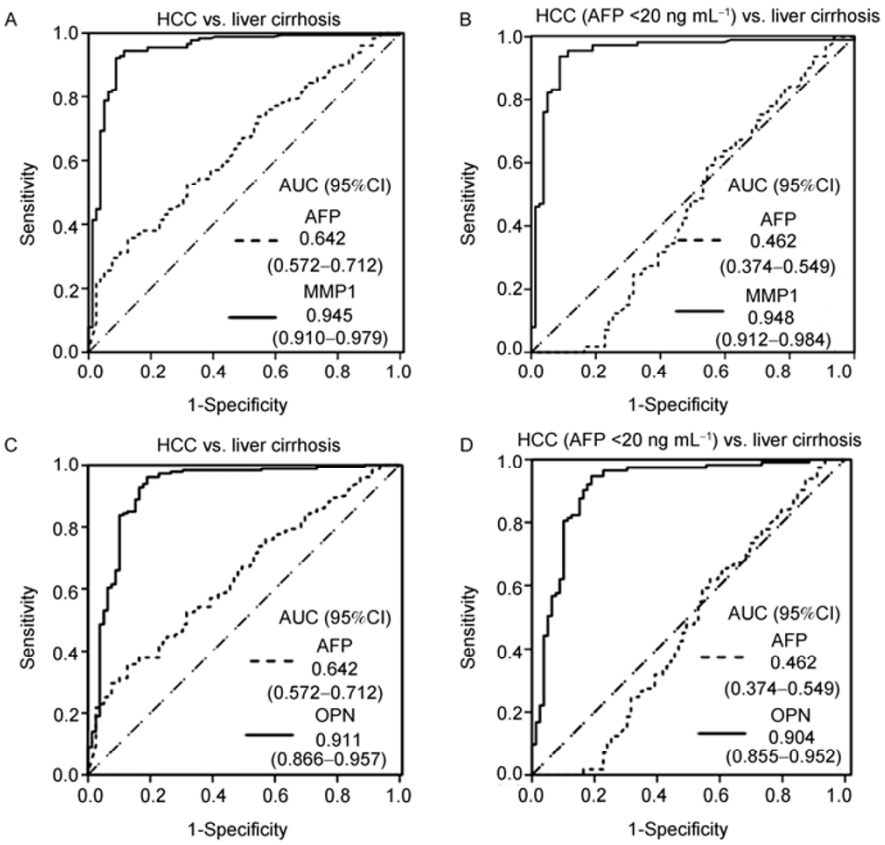

Figure 3 ROC evaluation of the ability of MMP1 and OPN to differentiate HCC patients from liver cirrhosis patients. The AUC is shown with 95% CI. A, MMP1 for HCC versus liver cirrhosis. B, MMP1 for HCC with AFP levels less than 20 ng mL^{-1} versus liver cirrhosis. C, OPN for HCC versus liver cirrhosis. D, OPN for HCC with AFP levels less than 20 ng mL^{-1} versus liver cirrhosis. AFP is represented by the dashed line, and MMP1 or OPN is represented by the solid line. CI, confidence interval.

Table 2 Area under the ROC curve (AUC), sensitivity and specificity using clinical AFP and MMP1 and OPN optimal cutoff values[a)]

Protein biomarkers	Cutoff (ng mL^{-1})	AUC (95% CI)	Sensitivity (%) (95% CI)	Specificity (%) (95% CI)
HCC versus healthy controls				
AFP	20.00	0.761 (0.706–0.815)	36.31 (29.30–43.80)	100.00 (96.50–100.00)
MMP1	1.72	0.876 (0.822–0.929)	92.18 (87.20–95.70)	85.44 (77.10–91.60)
OPN	84.40	0.935 (0.906–0.964)	87.15 (81.30–91.70)	89.32 (81.70–94.50)
PSG9	15.44	0.600 (0.533–0.668)	40.78 (33.50–48.40)	78.64 (69.50–86.10)
HCC versus liver cirrhosis				
AFP	20.00	0.642 (0.572–0.712)	36.87 (29.80–44.40)	84.81 (75.00–91.90)
MMP1	1.72	0.945 (0.910–0.979)	91.62 (86.60–95.20)	91.25 (82.80–96.40)
OPN	84.40	0.911 (0.866–0.957)	87.15 (81.30–91.70)	85.00 (75.30–92.00)
HCC (AFP<20 ng mL^{-1}) versus liver cirrhosis				
MMP1	1.72	0.948 (0.912–0.984)	92.92 (86.50–96.90)	91.25 (82.80–96.64)
OPN	84.40	0.904 (0.855–0.952)	84.96 (77.00–91.00)	85.00 (75.30–92.00)

a) CI, confidence interval.

3 Discussion

Cancer secretome studies are a good alternative to body fluid proteomics. Cancer secretome studies can increase analysis specificity while minimizing the number of biomarkers tested, particularly with respect to high-abundance proteins found in plasma [21]. One approach for biomarker discovery is the analysis of CM from HCC cell lines [10]. This proteome is less complex than the full plasma proteome and thus may not adequately represent the pathological status of this malignant disease [12]. In the present study, primary organ cultures were successfully established from cancerous and surrounding noncancerous tissues from HCC patients in a serum-free primary culture system. Using one-dimensional SDS-PAGE separation coupled with LC-MS/MS to analyze CM, 1365 non-redundant proteins were ultimately identified, 84% of which were found in the HUPO liver and two plasma proteomic datasets. AFP and other proteins, including lamin B [22] and heat shock protein 70 [23], that were elevated in the plasma or tissues of HCC patients were identified among the CM proteins. These findings confirmed the accuracy of this HCC CM protein database and that the quantities of CM proteins extracted were sufficient for further analysis. Particular attention was placed on extracellular proteins from the CM. These proteins had the highest probability of being observed in the circulation and therefore serving as potential biomarkers. In general, this approach is a step forward in the discovery of potential tumor biomarkers secreted from the tumor microenvironment.

This HCC-derived "secretory/released" protein database presents new possibilities for understanding liver function in a comprehensive and exploratory manner. In this study, 26% of CM proteins had a molecular function of "hydrolase activity", including epoxide and alcohol hydrolase (Figure 1A). Moreover, the biological processes of CM proteins were enriched for the primary metabolism of proteins, amino acids and carbohydrates (Figure 1B). The molecular functions and biological processes of these CM proteins reflect the features of liver metabolism. The cellular components of the CM proteins most likely to enter the blood were classified as "extracellular proteins" and "plasma membrane proteins" (Figure 1C). These proteins are released by the classical secretion mechanism or are shed from the cell surface. Because of cell lysis and other non-classical secretion mechanisms, some intracellular proteins, such as actin, carbonyl reductase and glyceraldehyde-3-phosphate dehydrogenase, were present in the CM [8]. In addition, 217 differential proteins, such as the liver form of glycogen phosphorylase, may participate in specific HCC-related processes (Figure S1B).

A thorough study with extensive follow-up validation of the selected proteins was then performed using plasma samples from a relatively large cohort of patients with HCC and liver cirrhosis that were chronically infected with HBV. The selected proteins were involved in TGF-β signaling pathways. TGF-β is a critical regulator of cell renewal, differentiation, and survival and is commonly deregulated in HCC. The suppression of TGF-β is an early event during hepatocarcinogenesis. Hepatitis B virus X proteins shift hepatic TGF-β signaling from tumor suppression to oncogenesis in patients with chronic HBV infection [24]. Thus, plasma biomarker candidates associated with TGF-β will likely help detect HCC in the early stages. OPN has been reported to promote oncogenesis and tumor cell invasion in HCC [25]. In this study, OPN was demonstrated to be a potential biomarker for the detection of HCC. This finding is consistent with a previous report suggesting OPN was more sensitive than AFP as a biomarker, and was a marker of early-stage HCC [26]. In addition, up-regulated OPN has been shown to indicate a poor prognosis for HCC [27]. PSG9 belongs to the carcinoembryonic antigen family and is a highly glycosylated protein that is deregulated during colorectal carcinogenesis [28]. PSG1 regulates immunoreactions and has proangiogenic properties [29]. Furthermore, PSG9 interacts with AFP and is involved in the TGF-β signaling pathway, based on network analysis (Figure S3). Because most HCCs arise in the background of chronic cirrhosis, the plasma levels of PSG9 were elevated in patients with liver cirrhosis in this study (Figure 2), indicating that increased PSG9 levels can serve as a biomarker of liver cirrhosis.

MMPs are the principal mediators of alterations observed in the microenvironment during cancer progression [30]. Because MMPs regulate the important pathway of extracellular matrix turnover, high plasma levels of MMP7, MMP9, and TIMP1 in cirrhosis were observed, suggesting their involvement in the progressive accumulation of extracellular matrix proteins following liver injury [31]. The most important finding in the current study was the identification of MMP1 as a diagnostic biomarker for HCC. A previous meta-analysis demonstrated that MMP1 increased the risk of cancer metastasis [32]. Decock et al. [33] reported that plasma levels of MMP1 were lower in breast cancer patients than in healthy controls, and plasma levels of MMP1 were elevated in lung cancer patients [34]. We demonstrated that MMP1 plasma levels were significantly increased in HCC, indicating that MMP1 plasma levels vary among different types of human tumors. We confirmed the strong discriminatory power of MMP1 to distinguish HCC patients from healthy controls and cirrhosis patients. Overall, the sensitivity and specificity of MMP1 in differentiating HCC cases from liver cirrhosis cases were better than those of AFP and OPN. The best performance of MMP1 was obtained when distinguishing HCC patients from liver cirrhosis patients, and ROC analysis revealed its improved performance in discerning HCC with AFP levels less than 20 ng mL^{-1}. These findings indicate that MMP1 could be used as a biomarker to improve HCC diagnosis at an early stage. This is

an important first step in the evaluation of MMP1 as an HCC marker. However, this study had some limitations, such as not using samples of long term HBV infection and not detecting dynamic changes in long-term follow-up plasma before HCC diagnosis. Further analysis will be needed to validate this biomarker in large-scale cohorts of HBV carriers.

In conclusion, this study reported an HCC secretome database of 1365 CM proteins, which was established based on a serum-free primary culture model. The data indicate this HCC cancer-related protein database provides a resource of potential biomarkers for diagnosis of this malignant disease. Among these validated biomarker candidates, MMP1 and OPN, which have better AUCs than AFP, may be potential plasma biomarkers to improve HCC diagnosis.

We thank Dr. Feng Lin for data analysis, and Dr. An Ning for his attentive review of this manuscript. We also thank Dr. Fu Yan and his group (Institute of Computing Technology, Chinese Academy of Sciences) for their generous support with the valuable pFind 2.4 software. This work was supported by the National Natural Science Foundation of China (81172035, 30973388), the National Excellent Doctoral Dissertation of China (2007B68), the National High Technology Research and Development Program of China (2012AA020206), and the Basic Research Program of the Cancer Hospital, PUMC & CAMS (JK2009B08, LC2009B45). The authors declare no competing financial interest.

1 Jemal A, Bray F, Center M M, et al. Global cancer statistics. CA Cancer J Clin, 2011, 61: 69–90
2 Liang X, Bi S, Yang W, et al. Epidemiological serosurvey of hepatitis B in China—declining HBV prevalence due to hepatitis B vaccination. Vaccine, 2009, 27: 6550–6557
3 Portolani N, Coniglio A, Ghidoni S, et al. Early and late recurrence after liver resection for hepatocellular carcinoma: prognostic and therapeutic implications. Ann Surg, 2006, 243: 229–235
4 Shimada K, Sakamoto Y, Esaki M, et al. Analysis of prognostic factors affecting survival after initial recurrence and treatment efficacy for recurrence in patients undergoing potentially curative hepatectomy for hepatocellular carcinoma. Ann Surg Oncol, 2007, 14: 2337–2347
5 Mao Y, Yang H, Xu H, et al. Golgi protein 73 (GOLPH2) is a valuable serum marker for hepatocellular carcinoma. Gut, 2010, 59: 1687–1693
6 States D J, Omenn G S, Blackwell T W, et al. Challenges in deriving high-confidence protein identifications from data gathered by a HUPO plasma proteome collaborative study. Nat Biotechnol, 2006, 24: 333–338
7 Meng R, Gormley M, Bhat V B, et al. Low abundance protein enrichment for discovery of candidate plasma protein biomarkers for early detection of breast cancer. J Proteomics, 2011, 75: 366–374
8 Makridakis M, Vlahou A. Secretome proteomics for discovery of cancer biomarkers. J Proteomics, 2010, 73: 2291–2305
9 Planque C, Kulasingam V, Smith C R, et al. Identification of five candidate lung cancer biomarkers by proteomics analysis of conditioned media of four lung cancer cell lines. Mol Cell Proteomics, 2009, 8: 2746–2758
10 Polisetty R V, Gupta M K, Nair S C, et al. Glioblastoma cell secretome: analysis of three glioblastoma cell lines reveal 148 non-redundant proteins. J Proteomics, 2011, 74: 1918–1925
11 Xiao T, Ying W, Li L, et al. An approach to studying lung cancer-related proteins in human blood. Mol Cell Proteomics, 2005, 4: 1480–1486
12 Zhang Y, Xu B, Liu Y, et al. The ovarian cancer-derived secretory/releasing proteome: a repertoire of tumor markers. Proteomics, 2012, 12: 1883–1891
13 Lechner J F, LaVeck M A. A serum-free method for culturing normal human bronchial epithelial cells at clonal density. J Tissue Cult Methods, 1985, 9: 43–48
14 Shen Z, Li P, Ni R J, et al. Label-free quantitative proteomics analysis of etiolated maize seedling leaves during greening. Mol Cell Proteomics, 2009, 8: 2443–2460
15 Sun H, Fang H, Chen T, et al. GOFFA: gene ontology for functional analysis—a FDA gene ontology tool for analysis of genomic and proteomic data. BMC Bioinformatics, 2006, 7(Suppl 2): S23
16 Wu S, Wan P, Li J, et al. Multi-modality of pI distribution in whole proteome. Proteomics, 2006, 6: 449–455
17 Omenn G S, States D J, Adamski M, et al. Overview of the HUPO Plasma Proteome Project: results from the pilot phase with 35 collaborating laboratories and multiple analytical groups, generating a core dataset of 3020 proteins and a publicly-available database. Proteomics, 2005, 5: 3226–3245
18 Anderson N L, Polanski M, Pieper R, et al. The human plasma proteome: a nonredundant list developed by combination of four separate sources. Mol Cell Proteomics, 2004, 3: 311–326
19 Chinese Human Liver Proteome Profiling Consortium. First insight into the human liver proteome from PROTEOME(SKY)- LIVER(Hu) 1.0, a publicly available database. J Proteome Res, 2010, 9: 79–94
20 Lin L, Amin R, Gallicano G I, et al. The STAT3 inhibitor NSC 74859 is effective in hepatocellular cancers with disrupted TGF-beta signaling. Oncogene, 2009, 28: 961–972
21 Pavlou M P, Diamandis E P. The cancer cell secretome: a good source for discovering biomarkers? J Proteomics, 2010, 73: 1896–1906
22 Sun S, Xu M Z, Poon R T, et al. Circulating Lamin B1 (LMNB1) biomarker detects early stages of liver cancer in patients. J Proteome Res, 2010, 9: 70–78
23 Luk J M, Lam C T, Siu A F, et al. Proteomic profiling of hepatocellular carcinoma in Chinese cohort reveals heat-shock proteins (Hsp27, Hsp70, GRP78) up-regulation and their associated prognostic values. Proteomics, 2006, 6: 1049–1057
24 Murata M, Matsuzaki K, Yoshida K, et al. Hepatitis B virus X protein shifts human hepatic transforming growth factor (TGF)-beta signaling from tumor suppression to oncogenesis in early chronic hepatitis B. Hepatology, 2009, 49: 1203–1217
25 Takafuji V, Forgues M, Unsworth E, et al. An osteopontin fragment is essential for tumor cell invasion in hepatocellular carcinoma. Oncogene, 2007, 26: 6361–6371
26 Shang S, Plymoth A, Ge S, et al. Identification of osteopontin as a novel marker for early hepatocellular carcinoma. Hepatology, 2012, 55: 483–490
27 Sun J, Xu H M, Zhou H J, et al. The prognostic significance of preoperative plasma levels of osteopontin in patients with TNM stage-I of hepatocellular carcinoma. J Cancer Res Clin Oncol, 2010, 136: 1–7
28 Salahshor S, Goncalves J, Chetty R, et al. Differential gene expression profile reveals deregulation of pregnancy specific beta1 glycoprotein 9 early during colorectal carcinogenesis. BMC Cancer, 2005, 5: 66
29 Lisboa F A, Warren J, Sulkowski G, et al. Pregnancy-specific glycoprotein 1 induces endothelial tubulogenesis through interaction with cell surface proteoglycans. J Biol Chem, 2011, 286: 7577–7586
30 Kessenbrock K, Plaks V, Werb Z. Matrix metalloproteinases: regulators of the tumor microenvironment. Cell, 2010, 141: 52–67
31 Asselah T, Bieche I, Laurendeau I, et al. Liver gene expression signature of mild fibrosis in patients with chronic hepatitis C. Gastroenterology, 2005, 129: 2064–2075
32 Liu D, Guo H, Li Y, et al. Association between polymorphisms in the promoter regions of matrix metalloproteinases (MMPs) and risk of cancer metastasis: a meta-analysis. PLoS ONE, 2012, 7: e31251
33 Decock J, Hendrickx W, Vanleeuw U, et al. Plasma MMP1 and MMP8 expression in breast cancer: protective role of MMP8 against

lymph node metastasis. BMC Cancer, 2008, 8: 77
34 Li M, Xiao T, Zhang Y, et al. Prognostic significance of matrix metalloproteinase-1 levels in peripheral plasma and tumour tissues of lung cancer patients. Lung Cancer, 2010, 69: 341–347

Supporting Information

Table S1 Clinicopathological characteristics of six paired samples from HCC patients

Figure S1 Proteins identified in the CM from six paired HCC tissues. A, Total number of proteins identified from the CM of paired cancerous (C) and surrounding noncancerous (N) tissues. B, The Venn diagram represented the comparison of HCC CM proteins with the published human proteomes.

Figure S2 Molecular weight (A) and pI (B) distribution of the identified CM proteins.

Figure S3 The interaction sub-network of the CM proteins involved in TGF-β signaling pathways. The proteins AFP, MMP1, MMP7, MMP9, OPN (SPP1), PSG9, and TIMP1 were subsequently validated as potential HCC biomarkers.

File S1 The information of protein identifications in HCC CM

The supporting information is available online at life.scichina.com and www.springerlink.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.